

Минобрнауки России

Федеральное государственное бюджетное образовательное учреждение
высшего образования

«Оренбургский государственный университет»

Кафедра информатики

РАБОЧАЯ ПРОГРАММА

ДИСЦИПЛИНЫ

«Б.1.В.ДВ.4.1 Статистические методы и модели в информационных системах»

Уровень высшего образования

БАКАЛАВРИАТ

Направление подготовки

09.03.02 Информационные системы и технологии
(код и наименование направления подготовки)

Общий профиль

(наименование направленности (профиля) образовательной программы)

Тип образовательной программы

Программа академического бакалавриата

Квалификация

Бакалавр

Форма обучения

Очная

Год набора 2016

Рабочая программа рассмотрена и утверждена на заседании кафедры

Кафедра информатики

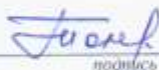
наименование кафедры

протокол № 5 от "22" 01 2016г.

Заведующий кафедрой

Кафедра информатики

наименование кафедры



подпись

М.А. Токарева

расшифровка подписи

Исполнители:

старший преподаватель кафедры информатики

должность



подпись

О.В. Юсупова

расшифровка подписи

должность

подпись

расшифровка подписи

СОГЛАСОВАНО:

Председатель методической комиссии по направлению подготовки

09.03.02 Информационные системы и технологии

код наименование



личная подпись

М.А. Токарева

расшифровка подписи

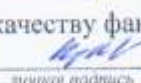
Заведующий отделом комплектования научной библиотеки


личная подпись

Н.Н. Грицай

расшифровка подписи

Уполномоченный по качеству факультета


личная подпись

И.В. Крючкова

расшифровка подписи

№ регистрации _____

© Юсупова О.В., 2016

© ОГУ, 2016

1 Цели и задачи освоения дисциплины

Цель (цели) освоения дисциплины: формирование знаний в области применения статистических методов и моделей, используемых для оценки эффективности информационных систем.

Задачи:

- сформировать целостное представление об основных методах статистического анализа;
- научить использовать существующие инструментальный, программное обеспечение и новейшие разработки в области статистического анализа оценки эффективности информационных систем;
- сформировать навыки применения различных статистических моделей, описывающих различные процессы жизненного цикла информационных систем;
- привить навыки проведения статистического анализа с использованием статистических пакетов программ.

2 Место дисциплины в структуре образовательной программы

Дисциплина относится к дисциплинам (модулям) по выбору вариативной части блока 1 «Дисциплины (модули)»

Пререквизиты дисциплины: *Б.1.Б.10.1 Алгебра и геометрия, Б.1.Б.10.3 Теория вероятностей и математическая статистика, Б.1.В.ОД.1 Объектно-ориентированное программирование*

Постреквизиты дисциплины: *Отсутствуют*

3 Требования к результатам обучения по дисциплине

Процесс изучения дисциплины направлен на формирование следующих результатов обучения

Планируемые результаты обучения по дисциплине, характеризующие этапы формирования компетенций	Формируемые компетенции
<u>Знать:</u> – основные средства реализации информационных технологий (методические, информационные, математические, алгоритмические, технические и программные); – функциональные возможности, ограничения и различия современных прикладных и методо-ориентированных пакетов программ для использования в профессиональной деятельности основ статистических методов. <u>Уметь:</u> – разрабатывать средства реализации информационных технологий (методические, информационные, математические, алгоритмические, технические и программные); – обрабатывать статистические данные в современных специализированных статистических пакетах программ. <u>Владеть:</u> – способностью разрабатывать средства реализации информационных технологий (методические, информационные, математические, алгоритмические, технические и программные); – навыками аналитической обработки материалов с применением возможностей современных статистических пакетов программ.	ПК-12 способностью разрабатывать средства реализации информационных технологий (методические, информационные, математические, алгоритмические, технические и программные)
<u>Знать:</u> – основные способы сбора, анализа научно-технической информации, отечественного и зарубежного опыта по тематике исследования;	ПК-22 способностью проводить сбор, анализ научно-

Планируемые результаты обучения по дисциплине, характеризующие этапы формирования компетенций	Формируемые компетенции
– основы функционирования систем сбора и обработки статистической информации с использованием современных программных статистических комплексов. Уметь: – проводить сбор, анализ научно-технической информации, отечественного и зарубежного опыта по тематике исследования; – представлять результаты статистического исследования с помощью наглядных презентаций с применением возможностей мультимедиа. Владеть: – способностью проводить сбор, анализ научно-технической информации, отечественного и зарубежного опыта по тематике исследования; – способностью изучения и анализа необходимой информации, технических данных, показателей и результатов работы, их обобщения и систематизации, проводить необходимые расчеты с использованием современных технических средств.	технической информации, отечественного и зарубежного опыта по тематике исследования
Знать: – области применения статистических методов и моделей, используемых для оценки эффективности информационных систем; – различные статистические модели, описывающие процессы жизненного цикла информационных систем. Уметь: – обосновывать правильность выбранной модели, сопоставляя результаты экспериментальных данных и полученных решений; – строить и оценивать статистические модели, описывающие процессы жизненного цикла информационных систем. Владеть: – способностью обосновывать правильность выбранной модели, сопоставляя результаты экспериментальных данных и полученных решений; – выбирать оптимальные статистические методы для принятия обоснованных решений на всех этапах жизненного цикла информационных систем.	ПК-24 способностью обосновывать правильность выбранной модели, сопоставляя результаты экспериментальных данных и полученных решений
Знать: – математические методы обработки, анализа и синтеза результатов профессиональных исследований; – основные методики и инструментарий применения статистических методов для решения задач профессиональных исследований. Уметь: – использовать математические методы обработки, анализа и синтеза результатов профессиональных исследований; – применять методики и инструментарий статистических методов для оценки эффективности информационных систем. Владеть: – способностью использовать математические методы обработки, анализа и синтеза результатов профессиональных исследований; – навыками обоснования возможности и эффективности применения статистических методов при решении задач профессиональных исследований.	ПК-25 способностью использовать математические методы обработки, анализа и синтеза результатов профессиональных исследований

4 Структура и содержание дисциплины

4.1 Структура дисциплины

Общая трудоемкость дисциплины составляет 5 зачетных единиц (180 академических часов).

Вид работы	Трудоемкость, академических часов	
	6 семестр	всего
Общая трудоёмкость	180	180
Контактная работа:	54,5	54,5
Лекции (Л)	18	18
Лабораторные работы (ЛР)	34	34
Консультации	1	1
Индивидуальная работа и инновационные формы учебных занятий	1	1
Промежуточная аттестация (зачет, экзамен)	0,5	0,5
Самостоятельная работа: - выполнение курсовой работы (КР); - написание реферата (Р); - самоподготовка (проработка и повторение лекционного материала и материала учебников и учебных пособий); - подготовка к лабораторным занятиям; - подготовка к рубежному контролю.	125,5 +	125,5
Вид итогового контроля (зачет, экзамен, дифференцированный зачет)	экзамен	

Разделы дисциплины, изучаемые в 6 семестре

№ раздела	Наименование разделов	Количество часов				
		всего	аудиторная работа			внеауд. работа
			Л	ПЗ	ЛР	
1	Основы вероятностно-статистических методов	88	8		20	60
2	Статистические методы анализа	92	10		14	68
	Итого:	180	18		34	128
	Всего:	180	18		34	128

4.2 Содержание разделов дисциплины

1 Основы вероятностно-статистических методов

Предмет, метод, задачи статистической науки. Систематизация и представление статистических данных. Элементы теории вероятностей: случайные события, случайные величины, законы распределения. Биноминальное распределение. Распределение Пуассона. Экспоненциальное распределение. Нормальное распределение. Распределение хи-квадрат. Распределение Стюдента. Распределение Фишера. Статистические методы: оценка параметров, проверка статистических гипотез. Гипотеза о равенстве дисперсий. Гипотеза о равенстве средних. Гипотеза о виде распределения. Критерии значимости при нормальном распределении. Критерии значимости при биномиальном распределении. Критерии согласия.

2 Статистические методы анализа

Дисперсионный анализ. Корреляционный анализ. Регрессионный анализ: парная линейная регрессия, парная нелинейная регрессия, множественная регрессия. Анализ случайных процессов: характеристики случайных процессов, анализ временных рядов. Методы многомерного анализа: метод главных компонент, дискриминантный анализ, кластерный анализ.

4.3 Лабораторные работы

№ ЛР	№ раздела	Наименование лабораторных работ	Кол-во часов
1	1	Случайные события и случайные величины: примеры решения задач	2
2	1	Использование электронных таблиц MS Excel для построения распределений случайных величин и генерации случайных чисел: нормальное распределение	2
3	1	Использование электронных таблиц MS Excel для построения распределений случайных величин и генерации случайных чисел: экспоненциальное распределение	2
4	1	Использование электронных таблиц MS Excel для построения распределений случайных величин и генерации случайных чисел: биномиальное распределение	2
5	1	Использование электронных таблиц MS Excel для построения распределений случайных величин и генерации случайных чисел: распределение Пуассона	2
6	1	Использование электронных таблиц MS Excel для построения распределений случайных величин и генерации случайных чисел: квантили распределений	2
7	1	Статистические методы в среде электронных таблиц MS Excel: метод описательной статистики	2
8	1	Статистические методы в среде электронных таблиц MS Excel: проверка гипотезы о равенстве дисперсий	2
9	1	Статистические методы в среде электронных таблиц MS Excel: проверка гипотезы о равенстве средних	2
10	1	Статистические методы в среде электронных таблиц MS Excel: проверка гипотезы о виде распределения	2
11	2	Однофакторный дисперсионный анализ в среде электронных таблиц MS Excel	2
12	2	Корреляционный анализ в среде электронных таблиц MS Excel	2
13	2	Регрессионный анализ в среде электронных таблиц MS Excel: линейная модель парной регрессии	2
14	2	Построение модели множественной регрессии в среде электронных таблиц MS Excel	2
15, 16	2	Анализ временных рядов в среде электронных таблиц MS Excel	4
17	2	Методы многомерного анализа: метод главных компонент, дискриминантный анализ, кластерный анализ	2
		Итого:	34

4.4 Курсовая работа (6 семестр)

Целью курсовой работы по дисциплине «Статистические методы и модели в информационных системах» является закрепление и углубление навыков применения статистического анализа.

Примерные темы:

1. Методы математической статистики
2. Дисперсионный анализ
3. Парная регрессия
4. Множественная регрессия
5. Анализ временных рядов
6. Методы многомерной классификации

По каждой теме курсовой работы предлагается выполнить типовый расчет.

Тема «Методы математической статистики»

1. Используя генератор случайных чисел пакета Анализ данных процессора электронных таблиц Excel, смоделируйте генеральную совокупность из N нормально распределенных чисел с заданным математическим ожиданием m_0 и стандартным отклонением σ .

2. Сформируйте из этой совокупности случайным образом выборку объемом n значений (*Анализ данных / Выборка*).

3. Для полученной выборки постройте гистограмму частот и график выборочной функции распределения, разбив данные на заданное число интервалов k .

4. Найдите точечные оценки математического ожидания и дисперсии. Используя эти оценки, нанести кривую нормального распределения на гистограмму частот.

5. Постройте 95-процентный доверительный интервал для математического ожидания (с использованием таблиц квантилей).

6. Проверьте гипотезу о том, что математическое ожидание генеральной совокупности равно заданному значению m_0 .

7. Разделите выборку на две части (первые n_1 значений – первая часть, остальные – вторая), и проверьте гипотезу о равенстве средних (двумя способами:

1) предполагая, что дисперсия генеральной совокупности известна (задано стандартное отклонение σ);

2) предполагая, что дисперсия генеральной совокупности неизвестна. Вначале надо проверить гипотезу о равенстве дисперсий – *Анализ данных / Двухвыборочный F-тест для дисперсий*, и в зависимости от результатов проверки использовать *Двухвыборочный t-тест* для средних с одинаковыми или различными дисперсиями).

N	m_0	σ	n	k	n_1
600	-10	2,5	40	8	15

Тема «Дисперсионный анализ»

Проверьте влияние фактора A на признак X . В таблице представлены результаты n наблюдений x_{ij} ($i = 1, 2 \dots n$) на различных уровнях фактора A_j ($j = 1, 2 \dots k$). Расчет надо провести с использованием таблицы квантилей и в пакете *Анализ данных / Однофакторный дисперсионный анализ* MS Excel.

$A1$	$A2$	$A3$
1,53	1,87	1,87
1,43	1,75	1,56
1,76	1,98	1,52
1,52	1,88	1,67
1,68	1,80	1,59

Тема «Парная регрессия»

Для заданной выборки:

1. Постройте диаграмму рассеяния.

2. Найдите параметры парной линейной регрессионной модели.
3. Нанесите линию регрессии на диаграмму рассеяния.
4. Проверьте значимость построенной модели.
5. Вычислите коэффициент детерминации.

6. Используя инструмент *Линия тренда* MS Excel, опробуйте построение линейной и других вариантов регрессионных моделей: экспоненциальной, степенной, логарифмической, полиномиальных 2-4 степеней. Результат для каждой модели представьте графически: диаграмма рассеяния, линия регрессии, ее уравнение, коэффициент детерминации. Необходимо выбрать регрессию, наиболее адекватную опытным данным.

x	2,0	2,5	3,0	3,5	4,0	4,5	5,0	5,5
y	8,34	8,21	8,07	7,93	7,89	7,54	7,34	7,12

Тема «Множественная регрессия»

Найдите параметры множественной линейной регрессии. Проверьте значимость полученной модели и значимость каждого фактора. Найдите коэффициент детерминации. Расчет проверьте, используя инструмент *Регрессия* пакета *Анализ данных* MS Excel.

x_1	1	3	5	7	9	11	13	15	17	19	21	23	25	27	29	31	1	3	5	7
x_2	0,8	1	1,4	1,9	2,4	2,2	2,7	2,5	3	3,2	2,6	2,4	2	1,5	1,8	3	0,8	1	1,4	1,9
x_3	9	12	21	13	21	24	35	28	15	35	28	35	15	24	25	9	12	21	13	21
y	20	40	50	60	90	100	130	100	130	180	170	150	150	170	200	200	2	5	8	6

Тема «Анализ временных рядов»

Постройте график заданного временного ряда. Проведите трехточечное сглаживание скользящей средней. Используя экспоненциальное сглаживание, спрогнозируйте ожидаемое значение при $t = 14$. Постройте линейный тренд и проверьте его значимость. Если тренд значим, найдите остатки и построьте для них авторегрессию первого порядка, при незначимом тренде построьте авторегрессию первого порядка для исходных данных. Проверьте значимость авторегрессии. Расчет необходимо провести, используя возможности MS Excel.

t	1	2	3	4	5	6	7	8	9	10	11	12	13
x	8	12	9	14	11	16	17	16	21	20	24	22	25

Тема «Методы многомерной классификации»

В таблице представлены данные о трех характеристиках 11 объектов, разделенных на 3 класса. Используя метод дискриминантного анализа, оцените, к какому классу относится 12-ый объект? Предполагая, что информация о разделении на классы отсутствует. Опробуйте несколько различных методов измерения расстояния между объектами и между кластерами. Сравните результаты, полученные двумя способами: с использованием дискриминантного и кластерного анализа.

x_1	2,1	2,0	1,8	1,5	1,7	1,6	1,5	1,3	1,5	1,2	1,2	1,6
x_2	4,7	4,8	4,7	4,9	4,8	4,9	5,3	5,2	5,4	5,1	5,1	4,8
x_3	6,9	6,9	6,7	6,6	6,4	6,3	6,0	6,1	6,2	6,2	6,0	6,6
класс	1	1	1	2	2	2	3	3	3	3	3	

5 Учебно-методическое обеспечение дисциплины

5.1 Основная литература

1. Гмурман, В.Е. Теория вероятностей и математическая статистика [Текст]: учебное пособие для бакалавров / В.Е. Гмурман. - 12-е изд. - Москва : Юрайт, 2014. - 479 с.
2. Рябченко, Н.В. Статистический анализ с применением программных средств: учебное пособие / Н.В. Рябченко, Е.В. Ларькина, И.И. Никитченко; Государственное казённое образовательное учреждение высшего профессионального образования «Российская таможенная академия» Владивостокский филиал, Кафедра административного и таможенного права. - Владивосток : Российская таможенная академия, Владивостокский филиал, 2015. - 112 с.: табл., схем. - Библиогр. в кн. - ISBN 978-5-9590-0634-1 ; То же [Электронный ресурс]. – Режим доступа:

<http://biblioclub.ru/index.php?page=book&id=438362>

3. Цыпин, А. П. Статистика в табличном редакторе Microsoft Excel [Электронный ресурс] : лабораторный практикум: учебное пособие для студентов, обучающихся по программам высшего образования по направлению подготовки 38.03.01 Экономика / А. П. Цыпин, Л. Р. Фаизова; М-во образования и науки Рос. Федерации, Федер. гос. бюджет. образоват. учреждение высш. образования "Оренбург. гос. ун-т". - Электрон. текстовые дан. (1 файл: 4.94 Мб). - Оренбург : ОГУ, 2016. - 288 с. – Режим доступа: http://artlib.osu.ru/web/books/metod_all/9988_20160329.pdf

5.2 Дополнительная литература

1. Костин, В.Н. Статистические методы и модели [Электронный ресурс]: учеб. пособие для вузов / В.Н. Костин, Н. А. Тишина; М-во образования Рос. Федерации, Гос. образоват. учреждение высш. проф. образования "Оренбург. гос. ун-т". - Оренбург : ГОУ ОГУ. - 2004. - 138 с- Загл. с тит. экрана. Издание на др. носителе: Статистические методы и модели [Текст] : учеб. пособие для вузов / В.Н. Костин, Н.А. Тишина; М-во образования Рос. Федерации, Гос. образоват. учреждение высш. проф. образования "Оренбург. гос. ун-т". - Оренбург : ГОУ ОГУ. - 2004. - 138 с. - Режим доступа: http://artlib.osu.ru/web/books/metod_all/516_20110701.pdf

2. Донскова, Л.И. Статистика: теория и практика: учебное пособие / Л.И. Донскова; Министерство образования и науки Российской Федерации, Негосударственное образовательное учреждение высшего профессионального образования «Западно-Сибирский Институт Финансов и Права». - Нижневартовск : Нижневартовский гуманитарный университет, 2012. - 275 с.: ил., табл. - Библиогр.: с. 256-257. - ISBN 978-5-94301-351-5; То же [Электронный ресурс]. – Режим доступа: <http://biblioclub.ru/index.php?page=book&id=429831>

3. Гмурман, В.Е. Руководство к решению задач по теории вероятностей и математической статистике [Текст]: учебное пособие для прикладного бакалавриата / В.Е. Гмурман.- 11-е изд., перераб. и доп. - Москва : Юрайт, 2016. - 404 с.

5.3 Периодические издания

1. Информатика и системы управления: журнал. - М.: Агентство "Роспечать", 2016
2. Информационные технологии в проектировании и производстве: журнал. - М.: Агентство "Роспечать", 2016
3. Открытые системы. СУБД: журнал. - М.: Агентство "Роспечать", 2016
4. Программные продукты и системы: журнал. - М.: Агентство "Роспечать", 2016.

5.4 Интернет-ресурсы

1. <http://www.gks.ru/> – Федеральная служба государственной статистики (официальная статистическая информация)
2. <http://orenstat.gks.ru/> – Территориальный орган Федеральной службы государственной статистики по Оренбургской области
3. <https://issek.hse.ru/> - Институт статистических исследований и экономики знаний
4. <http://www.edu.ru> – Федеральный портал «Российское образование»
5. <http://statsoft.ru/> – мультимедийный портал Академии Анализа Данных
6. <http://citforum.ru/> – IT-портал, содержащий технические материалы по всем областям компьютерных технологий, а также материалы аналитического, обзорного или новостного характера, посвященные корпоративным применениям информационных технологий
7. <http://fcior.edu.ru/> – Федеральный центр информационно-образовательных ресурсов
8. <http://edu-top.ru/katalog/> – Каталог образовательных ресурсов сети Интернет
9. <https://e.lanbook.com/books/1993> - «ЭСБ издательства «Лань»», каталог курсов «Автоматизированные системы и информатика»
10. <https://rucont.ru/collections/5610> - «ЭСБ Руконт», каталог курсов «Информатика и вычислительная техника»

5.5 Программное обеспечение, профессиональные базы данных и информационные справочные системы современных информационных технологий

1. Операционная система Microsoft Windows
2. Пакет настольных приложений Microsoft Office (Word, Excel, PowerPoint, OneNote, Outlook, Publisher, Access)
3. Пакет офисных приложений МойОфис Стандартный (МойОфис Текст, МойОфис Таблица, МойОфис Презентация, МойОфис Почта)
4. Open Office/LibreOffice - свободный офисный пакет программ, включающий в себя текстовый и табличный редакторы, редактор презентаций и другие офисные приложения
5. ПО для работы с файлами PDF Adobe Reader. Разработчик: Adobe Systems. Бесплатно после принятия лицензионного соглашения на ПО Adobe. Режим доступа: <https://get.adobe.com/ru/reader/>
6. Автоматизированная интерактивная система сетевого тестирования - АИССТ (зарегистрирована в РОСПАТЕНТ, Свидетельство о государственной регистрации программы для ЭВМ №2011610456, правообладатель – Оренбургский государственный университет), режим доступа - <http://aist.osu.ru>
7. Свободное распространяемое программное обеспечение для статистического анализа данных SalStat. Разработчик: Alan James Salmoni. Режим доступа: salstat.sourceforge.net
8. Свободное программное обеспечение для статистического анализа данных PSPP текущей версии. Тип лицензии: GNU GPL. Разработчик: GNU Project. Режим доступа: <http://www.gnu.org/software/pspp/>
9. Свободное распространяемое программное обеспечение для статистического анализа данных SOFA Statistics. Разработчик: Grant Paton-Simpson. Режим доступа: <http://www.sofastatistics.com/home.php>
10. ГАРАНТ Платформа F1 [Электронный ресурс]: справочно-правовая система. / Разработчик ООО НПП «ГАРАНТ-Сервис», 119992, Москва, Воробьевы горы, МГУ, [1990–2016]. – Режим доступа в сети ОГУ для установки системы: \\fileserver1\GarantClient\garant.exe
11. Большая советская энциклопедия [Электронный ресурс]: универсальная справочная энциклопедия международного уровня. – Режим доступа: <https://bigenc.ru/> в локальной сети ОГУ
12. Информационная система «Единое окно доступа к образовательным ресурсам» - Сайт, предоставляющий свободный доступ к каталогу образовательных интернет-ресурсов и полнотекстовой электронной учебно-методической библиотеке для общего и профессионального образования. [Электронный ресурс]. – Режим доступа: <http://window.edu.ru/window/catalog>
13. Springer [Электронный ресурс] : база данных научных книг, журналов, справочных материалов / компания Springer Customer Service Center GmbH. – Режим доступа: <https://link.springer.com/> , в локальной сети ОГУ
14. База данных публикаций в научных журналах и патентов Web Of Science. – Режим доступа: <http://apps.webofknowledge.com/>
15. Бесплатная база данных ГОСТ [Электронный ресурс] / ГОСТы Единой системы конструкторской документации, Единой системы программной документации, Единой системы стандартов на автоматизированные системы управления (АСУ), системы технической документации на АСУ, комплекса стандартов на автоматизированные системы, системы стандартов по базам данных и др. – Режим доступа: <https://docplan.ru>

6 Материально-техническое обеспечение дисциплины

Учебные аудитории для проведения занятий лекционного типа, семинарского типа, курсового проектирования, для проведения групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации.

Аудитории оснащены комплектами ученической мебели, техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Для проведения лабораторных занятий используется компьютерный класс, оснащенный компьютерной техникой подключенной к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду ОГУ.

Помещение для самостоятельной работы обучающихся оснащены компьютерной техникой подключенной к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду ОГУ.